Homework 2

February 12, 2022

José Carlos Muñoz

excersize 1)

we know that

$$x_1 = 2 \qquad\qquad x_2 = 3 \qquad\qquad \frac{\partial L}{\partial o} = 5$$

$$o = x_1 * x_2$$

To find $\frac{\partial L}{\partial x_1}$ and $\frac{\partial L}{\partial x_2}$ we use the Chain rule which gives us $\frac{\partial L}{\partial o}\frac{\partial o}{\partial x_1}$ and $\frac{\partial L}{\partial o}\frac{\partial o}{\partial x_2}$ respectively. It can be derived that $\frac{\partial o}{\partial x_1}$ and $\frac{\partial o}{\partial x_2}$ are x_2 and x_1 respectively
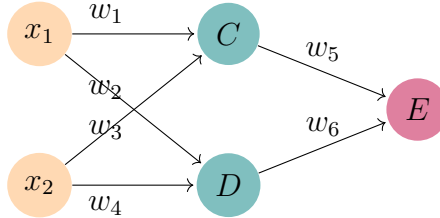
Therefore we can solve for both

$$\frac{\partial L}{\partial x_1} = \frac{\partial L}{\partial o}\frac{\partial o}{\partial x_1} \qquad\qquad\qquad \frac{\partial L}{\partial x_2} = \frac{\partial L}{\partial o}\frac{\partial o}{\partial x_2}$$
$$= 5 * x_2 \qquad\qquad\qquad\qquad = 5 * x_1$$
$$= 5 * 3 \qquad\qquad\qquad\qquad = 5 * 2$$
$$\frac{\partial L}{\partial x_1} = 15 \qquad\qquad\qquad\qquad \frac{\partial L}{\partial x_2} = 10$$

excersize 2)

The Neural Network is as follows.

We know that $w_1 = 0.1$, $w_2 = 0.5$, $w_3 = 0.4$, $w_4 = 0.3$, $w_5 = 0.2$, $w_6 = 0.6$. The Hidden Layer and Actrivation Layer,$y_h()$ and $y_o()$ respectively , both have the activation function of $y_n(z) = \frac{1}{1+e^{-z}}$. The Loss function is $L = \frac{1}{2}(y - \hat{y})^2$, where y is the expected value and $\hat{y}$ is the actual value

Our starting point is $\left(\begin{bmatrix} 0.82 \\ 0.23 \end{bmatrix} \quad 0\right)$. The Weights for each nodes are as follows

$$\vec{w_C} = \begin{bmatrix} w_1 \\ w_3 \end{bmatrix} \qquad\qquad \vec{w_D} = \begin{bmatrix} w_2 \\ w_4 \end{bmatrix} \qquad\qquad \vec{w_E} = \begin{bmatrix} w_5 \\ w_6 \end{bmatrix}$$

The Weights for the Hidden and Output layer as as follow

$$W_h = \begin{bmatrix} w_C & w_D \end{bmatrix} \qquad\qquad\qquad W_o = \begin{bmatrix} w_E \end{bmatrix}$$
$$= \begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \end{bmatrix} \qquad\qquad\qquad\qquad = \begin{bmatrix} w_5 \\ w_6 \end{bmatrix}$$

The z_C, z_D and z_E can be written as followed

$$z_C = w_1 * x_1 + w3 * x_2 \qquad z_D = w_2 * x_1 + w4 * x_2 \qquad z_E = w_5 * y_C + w_6 * y_D$$

For solving forward propogation we do the following steps

$$W_h^T * \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} z_C \\ z_D \end{bmatrix} \tag{1}$$

$$\begin{bmatrix} y_n(z_C) \\ y_n(z_D) \end{bmatrix} = \begin{bmatrix} y_C \\ y_D \end{bmatrix} \tag{2}$$

$$W_o^T * \begin{bmatrix} y_C \\ y_D \end{bmatrix} = z_E \tag{3}$$

$$y_E = y_n(z_E) \tag{4}$$

Plugging in the values we get this

$$\begin{bmatrix} 0.1 & 0.5 \\ 0.4 & 0.3 \end{bmatrix}^T * \begin{bmatrix} 0.82 \\ 0.23 \end{bmatrix} = \begin{bmatrix} 0.174 \\ 0.479 \end{bmatrix}$$

$$\begin{bmatrix} y_h(0.174) \\ y_h(0.479) \end{bmatrix} = \begin{bmatrix} 0.5433906 \\ 0.6175177 \end{bmatrix}$$

$$\begin{bmatrix} 0.2 \\ 0.6 \end{bmatrix}^T * \begin{bmatrix} 0.5433906 \\ 0.6175177 \end{bmatrix} = 0.47918874$$

$$y_o(0.47918874) = 0.617556289$$

The ending value, y_E, is 0.617556289 and the other values are as follows

$$y_C = 0.5433906 \qquad\qquad y_D = 0.6175177$$
$$z_C = 0.174 \qquad\qquad y_D = 0.479 \qquad\qquad z_E = 0.471988174$$

For Back propagation, the outer layer weights will derived as

$$\frac{\partial L}{\partial w_5} = \frac{\partial L}{\partial y_e} * \frac{\partial y_e}{\partial z_e} * \frac{\partial z_e}{\partial w_5} \qquad \frac{\partial L}{\partial w_6} = \frac{\partial L}{\partial y_e} * \frac{\partial y_e}{\partial z_e} * \frac{\partial z_e}{\partial w_6} \qquad \delta_o = \frac{\partial L}{\partial y_e} * \frac{\partial y_e}{\partial z_e}$$

$$\frac{\partial L}{\partial w_5} = \delta_o * \frac{\partial z_e}{\partial w_5} \qquad\qquad \frac{\partial L}{\partial w_6} = \delta_o * \frac{\partial z_e}{\partial w_6}$$

For the weights that connect to Node C, we find the rate of change as

$$\frac{\partial L}{\partial w_1} = \frac{\partial L}{\partial y_c} * \frac{\partial y_c}{\partial z_c} * \frac{\partial z_c}{\partial w_1} \qquad \frac{\partial L}{\partial w_3} = \frac{\partial L}{\partial y_c} * \frac{\partial y_c}{\partial z_c} * \frac{\partial z_c}{\partial w_3} \qquad \frac{\partial L}{\partial y_c} = \frac{\partial L}{\partial y_e} * \frac{\partial y_e}{\partial z_e} * \frac{\partial z_e}{\partial y_c}$$

$$= \delta_o * \frac{\partial z_e}{\partial y_c}$$

For the weights that connect to Node D, we find the rate of change as

$$\frac{\partial L}{\partial w_2} = \frac{\partial L}{\partial y_d} * \frac{\partial y_d}{\partial z_d} * \frac{\partial z_d}{\partial w_2} \qquad \frac{\partial L}{\partial w_4} = \frac{\partial L}{\partial y_d} * \frac{\partial y_d}{\partial z_d} * \frac{\partial z_d}{\partial w_4} \qquad \frac{\partial L}{\partial y_d} = \frac{\partial L}{\partial y_e} * \frac{\partial y_e}{\partial z_e} * \frac{\partial z_e}{\partial y_d}$$

$$= \delta_o * \frac{\partial z_e}{\partial y_d}$$

Derivatives for L, y are as follow

$$\frac{\partial L}{\partial y_e} = y - y_e \qquad\qquad \frac{\partial y_n}{\partial z_n} = y_n(z_n) * (1 - y_n(z_n))$$

Now lets solve for each weights
weights w_5 and w_6

$$\delta_o = (0 - y_E) * (y_n(z_E) * (1 - y_n(z_E)))$$
$$= 0.09032574$$

$$\frac{\partial L}{\partial w_5} = \delta_o * \frac{\partial z_e}{\partial w_5} \qquad\qquad \frac{\partial L}{\partial w_6} = \delta_o * \frac{\partial z_e}{\partial w_6}$$
$$= \delta_o * y_C \qquad\qquad\qquad\quad = \delta_o * y_D$$
$$= 0.09032574 * 0.5433906 \qquad\quad = 0.09032574 * 0.6175177$$
$$= 0.049082158 \qquad\qquad\qquad = 0.055777743$$

weights w_1 and w_3

$$\frac{\partial L}{\partial y_C} = \delta_o * \frac{\partial z_e}{\partial y_C}$$
$$= \delta_o * w_5$$
$$= 0.09032574 * 0.2$$
$$= 0.018065148$$

$$\frac{\partial L}{\partial w_1} = \frac{\partial L}{\partial y_C} * \frac{\partial y_C}{\partial z_C} * \frac{\partial z_C}{\partial w_1} \qquad\qquad \frac{\partial L}{\partial w_3} = \frac{\partial L}{\partial y_C} * \frac{\partial y_C}{\partial z_C} * \frac{\partial z_C}{\partial w_3}$$
$$= 0.018065148 * (y_n(z_C) * (1 - y_n(z_C))) * 0.82 \qquad = 0.018065148 * (y_n(z_C) * (1 - y_n(z_C))) * 0.23$$
$$= 0.018065148 * 0.248117255 * 0.82 \qquad\qquad = 0.018065148 * 0.248117255 * 0.23$$
$$= 0.00367546544 \qquad\qquad\qquad\qquad\quad = 0.00103092323$$

weights w_2 and w_4

$$\frac{\partial L}{\partial y_D} = \delta_o * \frac{\partial z_e}{\partial y_D}$$
$$= \delta_o * w_6$$
$$= 0.09032574 * 0.6$$
$$= 0.05419544$$

$$\frac{\partial L}{\partial w_2} = \frac{\partial L}{\partial y_D} * \frac{\partial y_D}{\partial z_D} * \frac{\partial z_D}{\partial w_2} \qquad\qquad \frac{\partial L}{\partial w_4} = \frac{\partial L}{\partial y_D} * \frac{\partial y_D}{\partial z_D} * \frac{\partial z_D}{\partial w_4}$$
$$= 0.05419544 * (y_n(z_D) * (1 - y_n(z_D))) * 0.82 \qquad = 0.05419544 * (y_n(z_D) * (1 - y_n(z_D))) * 0.23$$
$$= 0.05419544 * 0.23618959 * 0.82 \qquad\qquad = 0.05419544 * 0.23618959 * 0.23$$
$$= 0.01049632697 \qquad\qquad\qquad\qquad\quad = 0.00294409171$$

The Final $\frac{\partial L}{\partial w}$ is as follows

$$
\frac{\partial L}{\partial \vec{w}} =
\begin{bmatrix}
0.00367546544 \\
0.01049632697 \\
0.00103092323 \\
0.00294409171 \\
0.049082158 \\
0.055777743
\end{bmatrix}
$$

The learning rate ϵ is 0.7 so the new weights are as followed

$$
\vec{w'} = \vec{w} - \epsilon * \frac{\partial L}{\partial \vec{w}}
$$

$$
=
\begin{bmatrix}
0.1 \\
0.5 \\
0.4 \\
0.3 \\
0.2 \\
0.6
\end{bmatrix}
- 0.7 *
\begin{bmatrix}
0.00367546544 \\
0.01049632697 \\
0.00103092323 \\
0.00294409171 \\
0.049082158 \\
0.055777743
\end{bmatrix}
$$

$$
=
\begin{bmatrix}
0.09742717419 \\
0.49265257112 \\
0.39927835373 \\
0.2979391358 \\
0.1656424894 \\
0.560955799
\end{bmatrix}
$$